Journal of Geophysical Research

Supporting Information for

Inverse Estimation of an Annual Cycle of California's Nitrous Oxide Emissions

Seongeun Jeong[1], Sally Newman[2*], Jingsong Zhang[3], Arlyn E. Andrews[4], Laura Bianco[4,5], Ed Dlugokencky[4], Justin Bagley[1], Xinguang Cui[1], Chad Priest[3], Mixtli Campos-Pineda[3], Marc L. Fischer[1]

[1]Lawrence Berkeley National Lab, Berkeley, CA

[2]California Institute of Technology, Pasadena, CA

[3]University of California, Riverside, CA

[4]Earth System Research Laboratory, NOAA, Boulder, CO

[5]Cooperative Institute for Research In Environmental Sciences, Univ. of Colorado, Boulder, CO

*Now at Bay Area Air Quality Management District, San Francisco, CA

Contents of this file

References for Supplemental Information

Text S1. Description of Individual Probability Distributions

For the likelihood function without bias correction (see Text S2 for bias correction) we use

$$\boldsymbol{y}|\boldsymbol{\lambda}, \boldsymbol{\sigma_R}, \eta, \tau \sim N(\boldsymbol{K\lambda}, \boldsymbol{R}) \qquad \text{(S1)}$$

where N is the normal distribution with mean $\mathbf{K\lambda}$ ($n \times 1$) and covariance $\mathbf{R}$ (n × n). Note that $\mathbf{y}$ is conditionally independent of all other parameters given $\boldsymbol{\lambda}$, $\boldsymbol{\sigma}_\mathbf{R}$, η and τ.

In order to estimate parameter values with Bayesian inference, prior uncertainty needs to be specified. In the hierarchical model, we need to include prior uncertainty for the joint parameter set $\boldsymbol{\Theta}$ (see Equation 2 in the main text) using a series of distributions. The scaling factor $\boldsymbol{\lambda}$ is sampled from a truncated normal distribution (instead of a fixed value) as

$$\boldsymbol{\lambda} \sim N(\boldsymbol{\mu_\lambda}, \boldsymbol{\sigma}_\lambda) \quad \text{(S2)}$$

where $\boldsymbol{\mu_\lambda}$ itself is sampled from a truncated normal distribution (Michalak, 2008; Miller et al., 2014) with a mean of 1 and a standard deviation of 0.5 so that 68% of the samples are within 50 ~ 150% from the mean (Ganesan et al., 2014; Jeong et al., 2016, 2017). $\boldsymbol{\sigma_\lambda}$ is modeled using a half Cauchy distribution, which is one of the recommended distributions for model variances (Gelman et al., 2014; Gelman and Hill, 2007; Korner-Nievergelt et al., 2015). The hyper-parameterization ("hyper" meaning the upper level in the hierarchy) for $\boldsymbol{\sigma_\lambda}$ can formally be expressed as

$$\boldsymbol{\sigma_\lambda} \sim hCauchy\,(0,1) \quad \text{(S3)}$$

where *hCauchy* is the half-Cauchy distribution. $\boldsymbol{\sigma_\lambda}$ is sampled from a distribution with a heavy tail so that $\boldsymbol{\sigma_\lambda}$ can be optimized from a broad distribution (instead of being a fixed value such as 50% of the mean emission). Equation S3 suggests that if we generate random samples (large enough) from Equation S3 we get a median value close to 1.

For the model-measurement covariance matrix $\mathbf{R}$, we use an exponential covariance function (Jeong et al., 2016, 2017; Rasmussen and Williams, 2006), which can be written in general form as

$$R_{i,j} = \eta^2 \exp\left(-\frac{1}{\tau}|t_i - t_j|\right) + \delta_{i,j}\sigma_{R_s}^2 \quad \text{(S4)}$$

where η, τ, and σ_{R_s} are parameters that define the covariance function, t is the measurement time, and δ is the Kronecker delta function (value of 1 if i = j, otherwise zero). Note that here we use the L_1 norm (i.e., $|t_i - t_j|$) (Ganesan et al., 2014; Jeong et al., 2016, 2017). The subscript s in σ_{R_s} indicates that σ_R is estimated for each site (Jeong et al., 2013, 2016, 2017).

We model σ_{R_S} using the half Cauchy distribution as in $\boldsymbol{\sigma_\lambda}$ (Gelman et al., 2014; Gelman and Hill, 2007; Korner-Nievergelt et al., 2015). The scale parameter (in the hyper-parameter sense) for the half Cauchy distribution for σ_{R_S} can be calculated using a first-order approximation method (Jeong et al., 2012a, 2012b, 2013) and used as

$$\sigma_{R_s} \sim hCauchy(0, \sigma_{R_{p|s}}) \quad \text{(S5)}$$

where $\sigma_{R_{p|s}}$ is the first-order estimate for σ_{R_s} and includes errors from several sources (e.g., transport and background errors) combined in quadrature. In this study, $\sigma_{R_{p|s}}$ is derived by scaling the estimate for σ_{R_S} from the WGC site (Jeong et al., 2012b) in proportion to the mean measured mole fraction for each site relative to that of WGC.

For η, we use non-informative prior as

$$\eta \sim unif(0, L) \quad \text{(S6)}$$

where η is allowed to vary from 0 to *L* with an equal probability of 1/*L*. Since η is generally estimated to be smaller than $\sigma_{R_{p|s}}$ we use $\sigma_{R_{p|s}}$ as an upper limit for *L* for efficient sampling (Jeong et al., 2016).

We use the exponential distribution (Ganesan et al., 2014; Jeong et al., 2016, 2017) for τ as

$$\tau \sim \exp\left(\frac{1}{\tau_p}\right) \quad \text{(S7)}$$

where τ_p is the hyper parameter for τ, which is assumed to be 7 days (typical synoptic time scale for transport).

Text S2. Bias Correction

As described in Section 2.4 of the main text, we correct bias during the inversion process treating the bias correction through a stochastic parameter (i.e., **D** from Equation 1 in the main text). Following Jeong et al. (2017), we modify the likelihood function in Equation S1 to

$$\boldsymbol{y}|\boldsymbol{\lambda}, \boldsymbol{\sigma_R}, \eta, \tau, \boldsymbol{D} \sim N(\boldsymbol{K\lambda} + \boldsymbol{D}, \boldsymbol{R}) \qquad \text{(S8)}$$

where we sample from a normal distribution with a mean of $\boldsymbol{K\lambda} + \boldsymbol{D}$ instead of $\boldsymbol{K\lambda}$, reflecting stochastic bias corrections during the inversion. It is worth noting that using this stochastic method, the uncertainty associated with bias correction is fully propagated into the inversion system.

The stochastic parameter **D** is sampled as

$$\boldsymbol{D} \sim unif\ (-5, 5) \qquad \text{(S9)}$$

where we use a uniform prior to sample for **D**. As a hyper-parameter in the form of [a, b], we used [-5, 5] (in the units of ppb) based on the prior information obtained from a simple linear regression method. Following Xiang et al. (2013), we calculated the bias using the intercept of linear regression analyses of predicted and measured mole fractions and found the maximum bias was ~3 ppb. Based on this, we use a larger interval (i.e., [-5, 5]) than 3 ppb.

The posterior distribution of estimated **D** is summarized in Figure S6 where the bias varies by season and site. Note that as demonstrated in Jeong et al. (2017) and also as shown in Equation S8, the values of **D** represent the total bias that includes the background bias, measurement offset and other potential biases that are not captured by our transport model. Because the bias at each site is estimated as a single offset value for each month that includes the combination of the background bias, measurement offset and other uncaptured transport biases in an additive fashion, it is difficult to distinguish one bias source from another. However, we note that as reported in Xiang et al. (2013), the background curtain may be a significant source of bias. We also note that some of our sites in the Central Valley (ARV and STB) have calibration uncertainties that limit measurement accuracy to 1 ppb, within a factor of 2 of estimated bias for those sites. With respect to transport, we reinforce previous comments by Bagley et al. (2017) that some sites may not be well represented by the closest available measurements (e.g., wind profiler data), and potential undiagnosed errors may exist depending on the site. Finally, we acknowledge that seasonal variation in bias adds another layer of complexity to resolving the total bias into different sources. Although our stochastic method fully propagates biases through our inversion system, separation of bias sources requires further studies and deserves in-depth analysis in the future. The values for **D** estimated from the inverse optimization using Equation S8 are generally comparable to values obtained from the simple regression analysis. The low bias in the predictions estimated from the regression was noticeable at ARV (1.1 – 2.7 ppb), CIT (0.34 – 1.7 ppb), SBC (0.2 - 0.7 ppb) and STB (~1 ppb) during spring and summer. At the annual scale, the stochastic bias correction (i.e., Equation S8) inversion yielded slightly lower annual total N_2O emissions compared to those of the simple regression-based method (79 vs. 87 Gg N_2O/yr in median values). For the case based on the simple regression method, the calculated bias was corrected before the inversion.

Text S3. Convergence and Accuracy of MCMC Samples

The Markov chain Monte Carlo (MCMC) sampler based on the JAGS system (Just Another Gibbs Sampler, Plummer, 2003) generated 50000 samples from the posterior distribution using two chains after discarding the 2000 burn-in samples. To check MCMC representativeness of the posterior distribution, we use visual inspection of trace plots and probability density plots. We also calculate the Gelman-Rubin statistic (also known as "potential scale reduction factor") (Gelman and Rubin, 1992). The value of one for the Gelman-Rubin statistic suggests the chains are fully converged. Values that are significantly larger than one indicate lack of convergence.

Figure S10 shows the Gelman-Rubin statistics for the scaling factor ($\boldsymbol{\lambda}$) for the mid-month of each season. As shown in Figure S10, the Gelman-Rubin statistics for $\boldsymbol{\lambda}$ is close to one for all elements of the scaling factor vector, showing the chains converged well. Figure S11 shows the Gelman-Rubin statistics for $\boldsymbol{\sigma_R}$ (parameter for model-measurement mismatch uncertainty) for four sites in the Central Valley and SoCAB with all sites showing good convergence. Other parameters also showed similar results in the convergence test. The burn-in period (i.e., preliminary period for the parameter to move from the initial values to the posterior region) typically requires several hundreds to several thousand steps. In this study, we use 2000 steps for the burn-in period, and the trace and density plots show that the burn-in period has been successfully passed. As an example, the trace and density plots for $\boldsymbol{\sigma_R}$ at WGC and CIT are shown in Figure S12 where the two chains are smooth (without staying on the same value for a long time) and overlap each other suggesting convergence has been achieved.

To check the accuracy of the chain we estimate the standard error (SE) for the MCMC posterior estimate, which is calculated as the standard deviation (SD) of the chain divided by the square root of the effective sample size (ESS). ESS can be calculated by dividing the number of points in the chain by the autocorrelation time (Kass et al., 1998). Thus, SE represents the estimated standard deviation for the sample mean in the chain (Kruschke, 2015). Since SE is estimated on the scale of the parameter, it is useful to assess the accuracy of the posterior mean estimate for the parameter. For example, Figure S12 shows the SE values for $\boldsymbol{\sigma_R}$ at CIT and WGC as well as density plots. As shown in the figure, compared to the mean (and the 95% confidence interval), the SE values are small indicating the posterior mean values are estimated accurately.

Table S1. WRF-STILT boundary layer – land surface model combination* used for predicting N_2O mole fractions at each site for each month between June 2013 and May 2014.

	ARV	CIT & STR	SBC	WGC & STB‡
June	NA†	MYNN2-Noah	MYNN2- Noah	MYJ– LSM
July	NA	MYNN2- Noah	MYNN2- Noah	MYJ - LSM
August	NA	MYNN2- Noah	MYNN2- Noah	MYJ - LSM
September	NA	MYNN2- Noah	MYNN2- Noah	NA
October	MYNN2- Noah	MYNN2- Noah	MYNN2- Noah	MYNN2- Noah
November	YSU- Noah	MYNN2- Noah	YSU- Noah	MYNN2- Noah
December	YSU- Noah	MYNN2- Noah	YSU- Noah	MYNN2- Noah
January	YSU- Noah	MYNN2- Noah	YSU- Noah	MYNN2- Noah
February	MYNN2- Noah	MYNN2- Noah	MYNN2- Noah	MYNN2- Noah
March	MYNN2- Noah	MYNN2- Noah	MYNN2- Noah	MYNN2- Noah
April	MYNN2-LSM	MYNN2- Noah	MYNN2- Noah	MYNN2- LSM
May	MYNN2-LSM	MYNN2- Noah	MYNN2- Noah	MYNN2- LSM

*For each combination, WRF was run independently.
†"NA" indicates that observations were not available during that month.
‡For STB, the STILT run was performed only for April and May when the measurements were available.

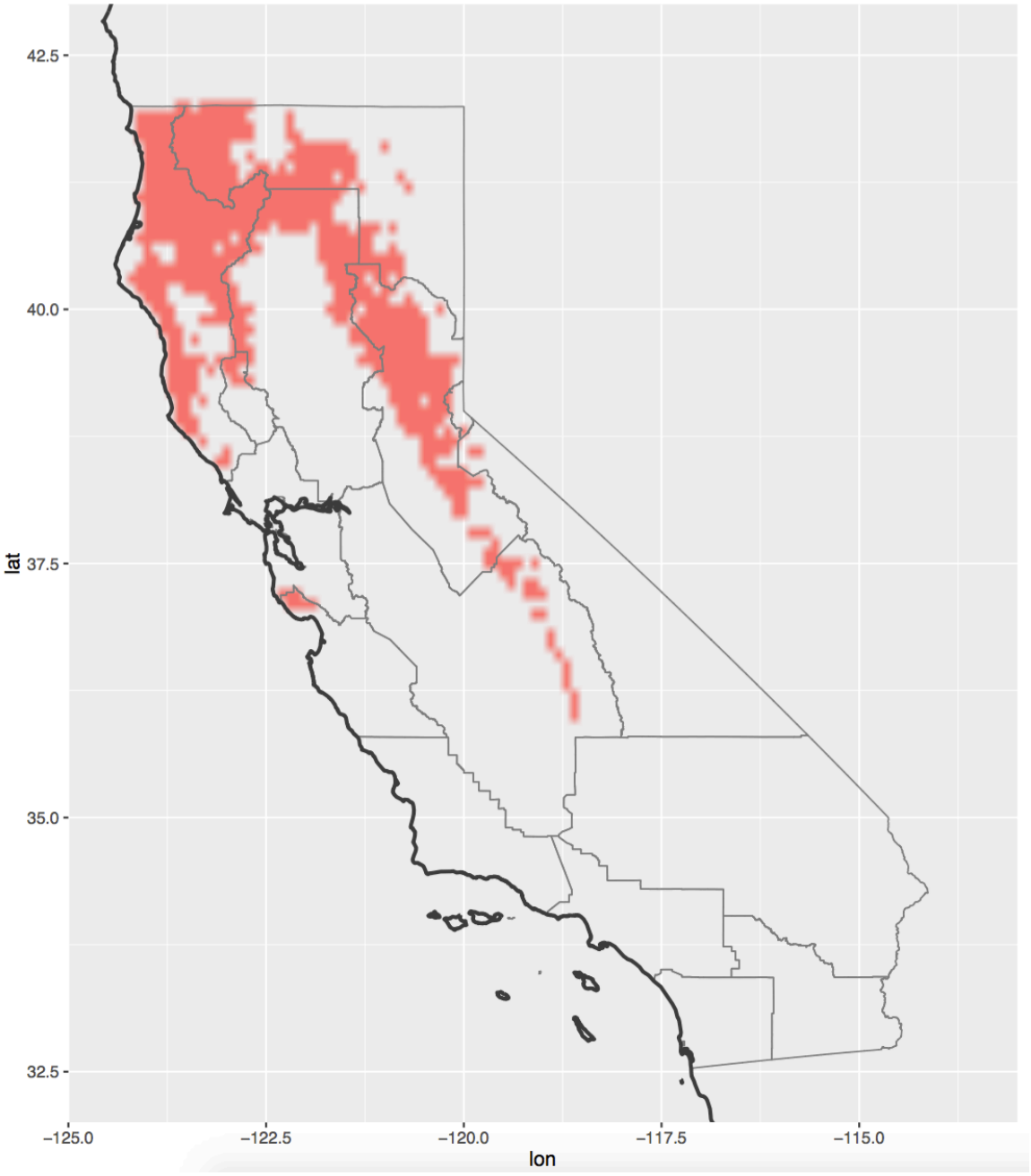

Figure S1. Identified natural forest pixels (~10 km) in California using the Moderate Resolution Imaging Spectroradiometer (MODIS) land cover type data product (MCD12Q1, year 2012).

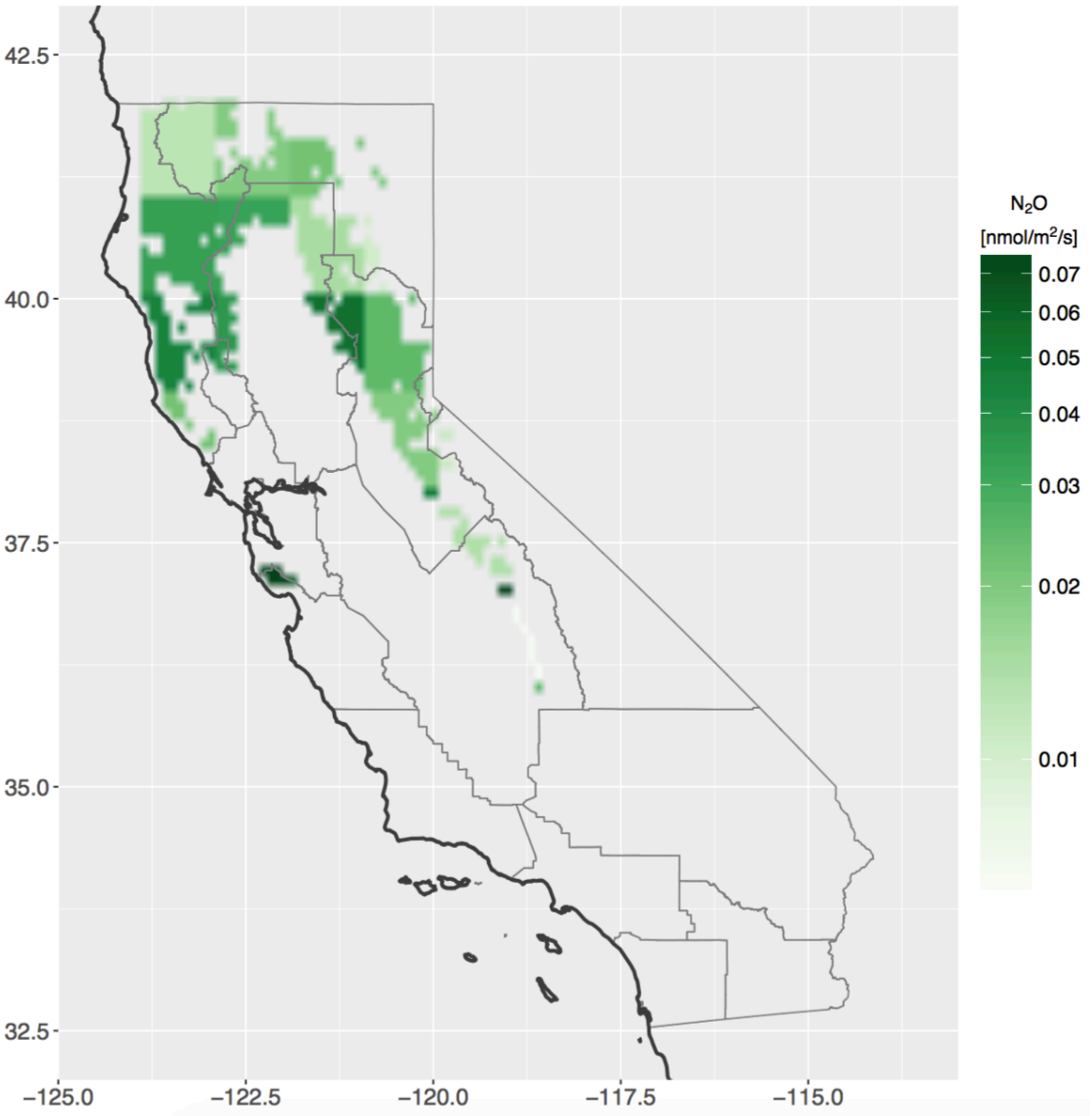

Figure S2. Derived natural forest N_2O emissions for California (nmol/m^2/s) based on the Global Emissions InitiAtive (GEIA) emission model.

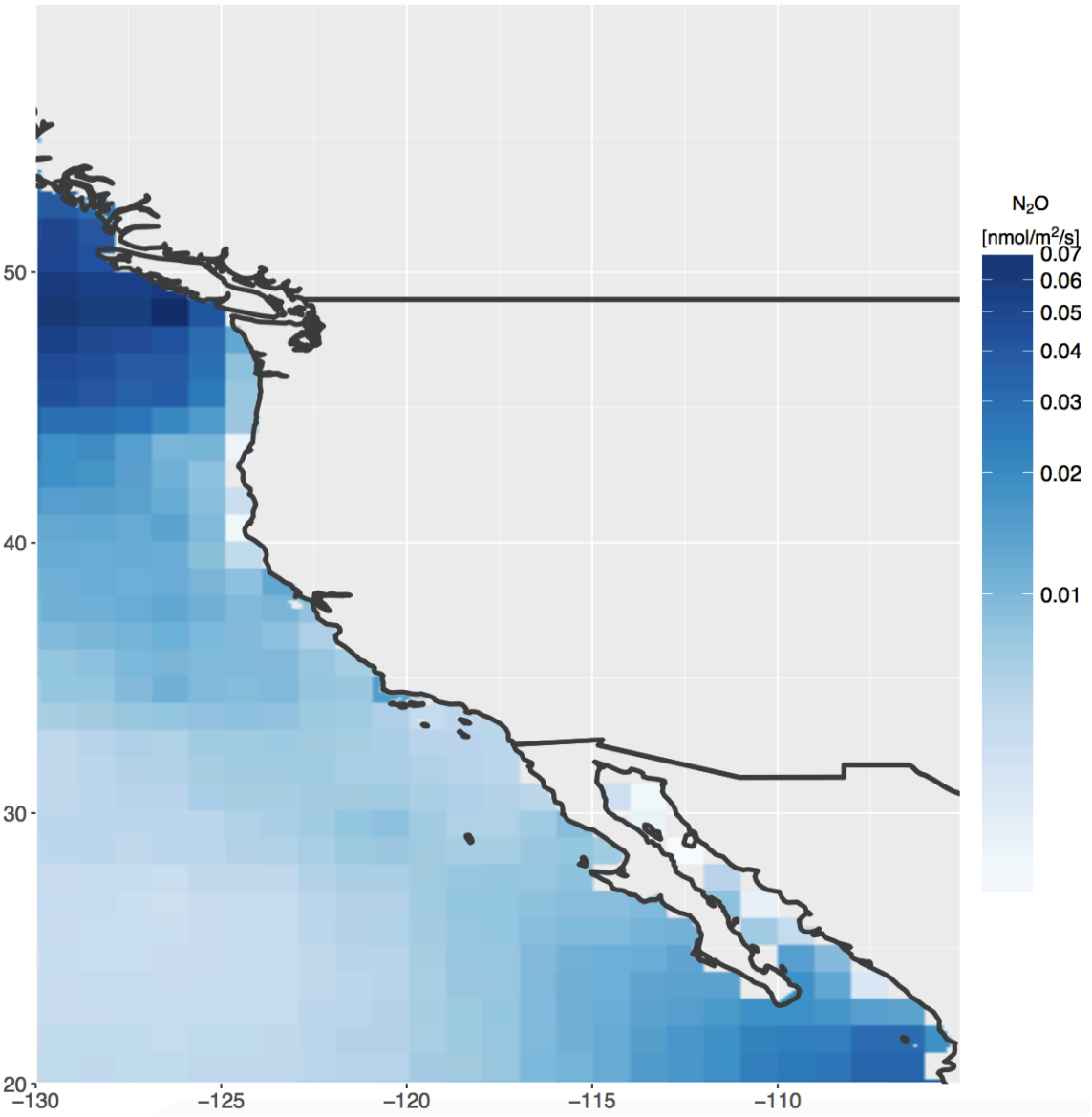

Figure S3. Regridded (at 10 km resolution) GEIA ocean N_2O emissions (nmol/m^2/s) for the entire modeling domain.

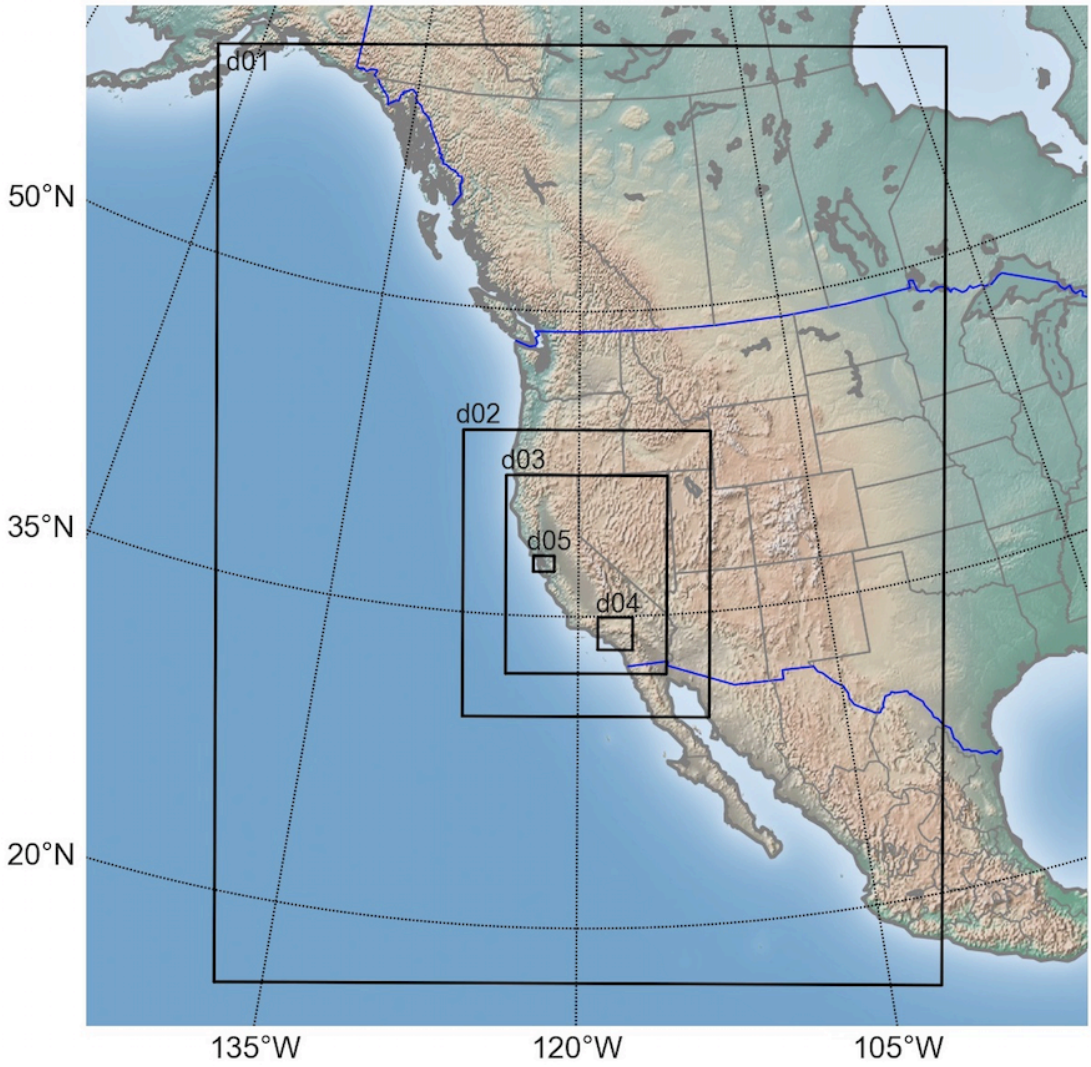

Figure S4. WRF modeling domain used in the inversion study. The actual inversion domain covers 130°W - 105°W and 20°N - 59.9°N for longitude and latitude, respectively. d01, d02, d03, d04 and d05 represent the WRF modeling domains of 36, 12, 4, 1.3, and 1.3 km, respectively.

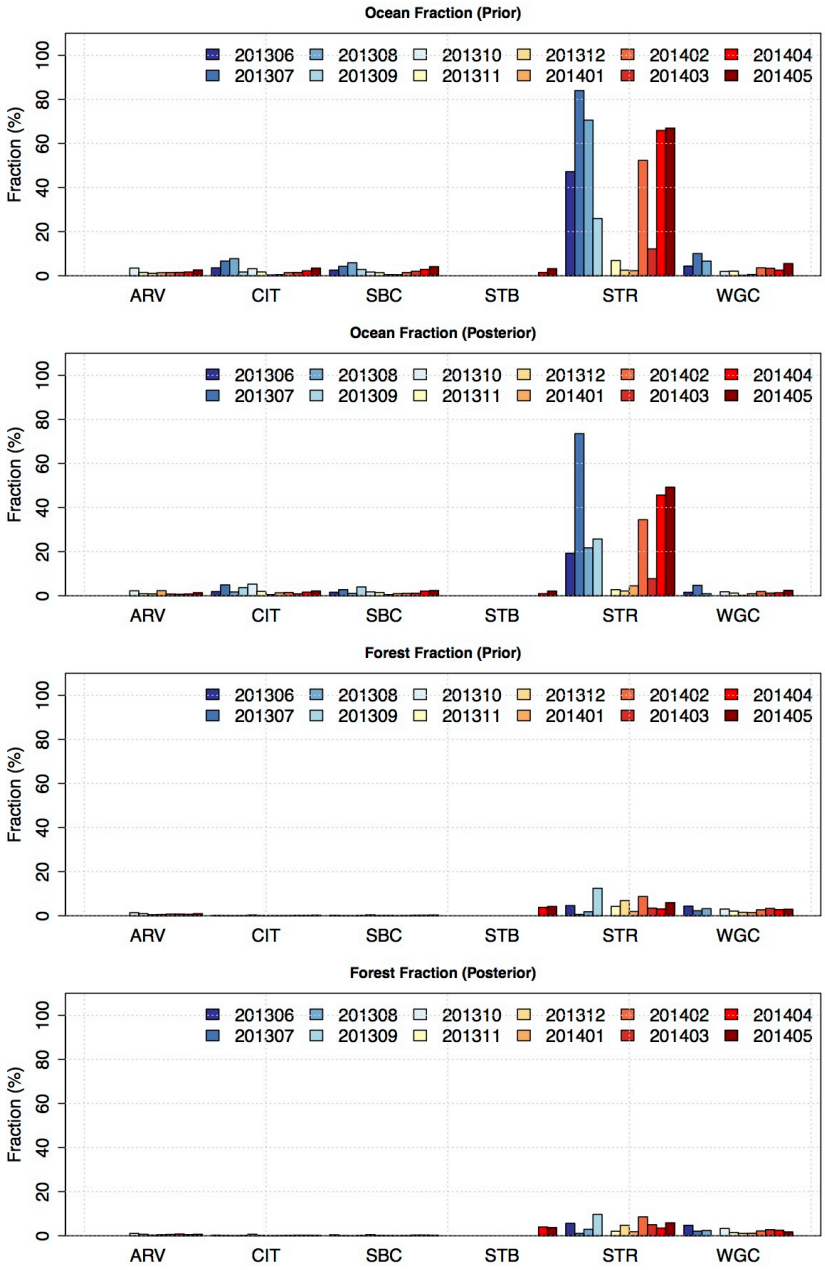

Figure S5. Predicted monthly mean mole fractions for ocean and forest as a fraction (%) of the total prediction by site and month. The top two panels show the ocean fraction before (prior) and after (posterior) inversion, respectively, and the bottom two panels show the forest fraction before and after inversion. Except for the coastal site, STR, the fraction for both forest and ocean is less than ~10%. Because most of the measurement sites are inland where the footprint influence of ocean emissions is weak, ocean signals are small except for STR. Note that although the total ocean emission from the prior emission map is comparable to the state total emission, ocean emissions are weighted by the weak footprint in the ocean (as compared to those on land), yielding generally less than 10% of the total mole fraction at most sites.

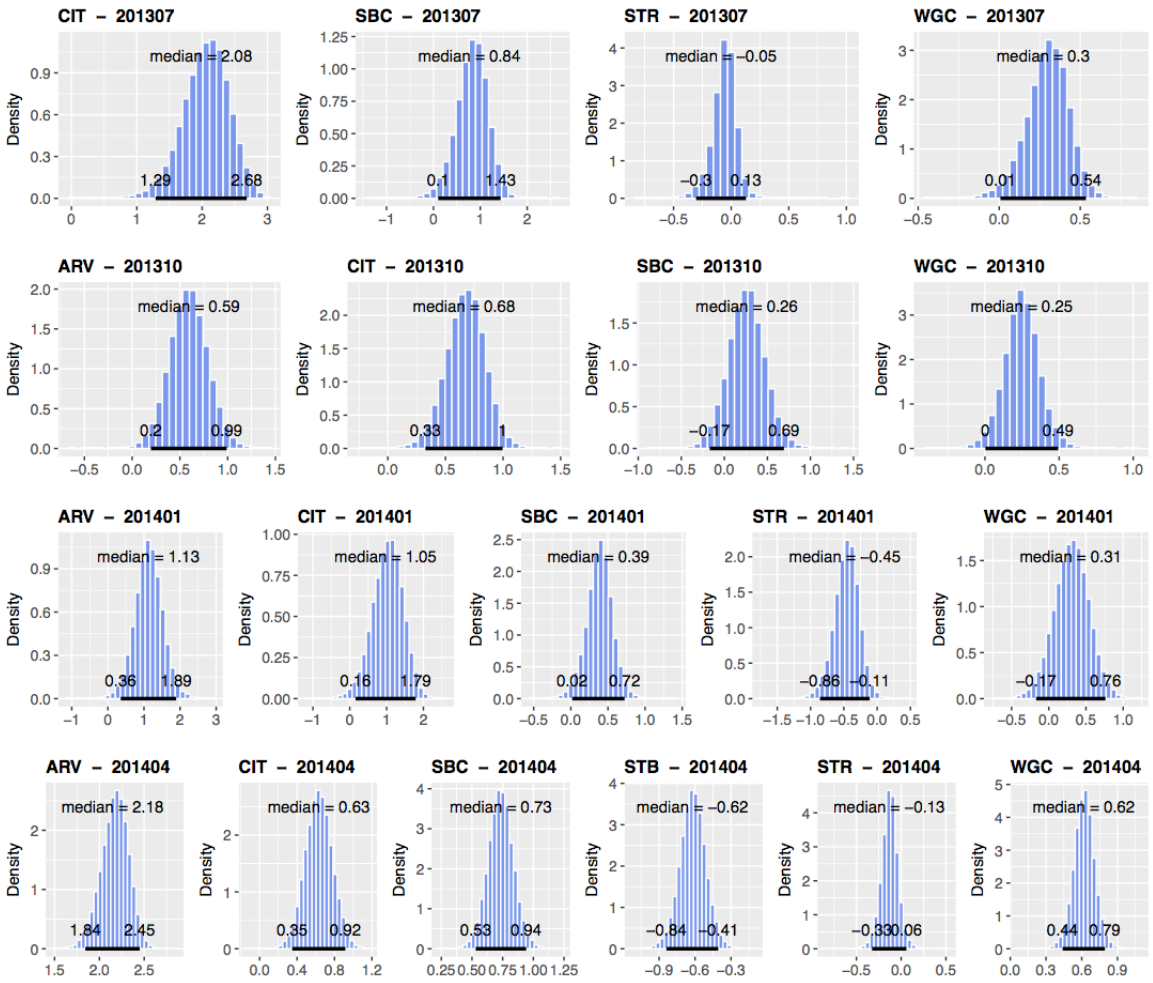

Figure S6. Summary for the posterior distribution of the bias correction parameter (i.e., **D** in Equation 1) for the mid-month of each season by site (shown for sites with measurements). The values at the bottom of each density plot represent the 95% confidence intervals.

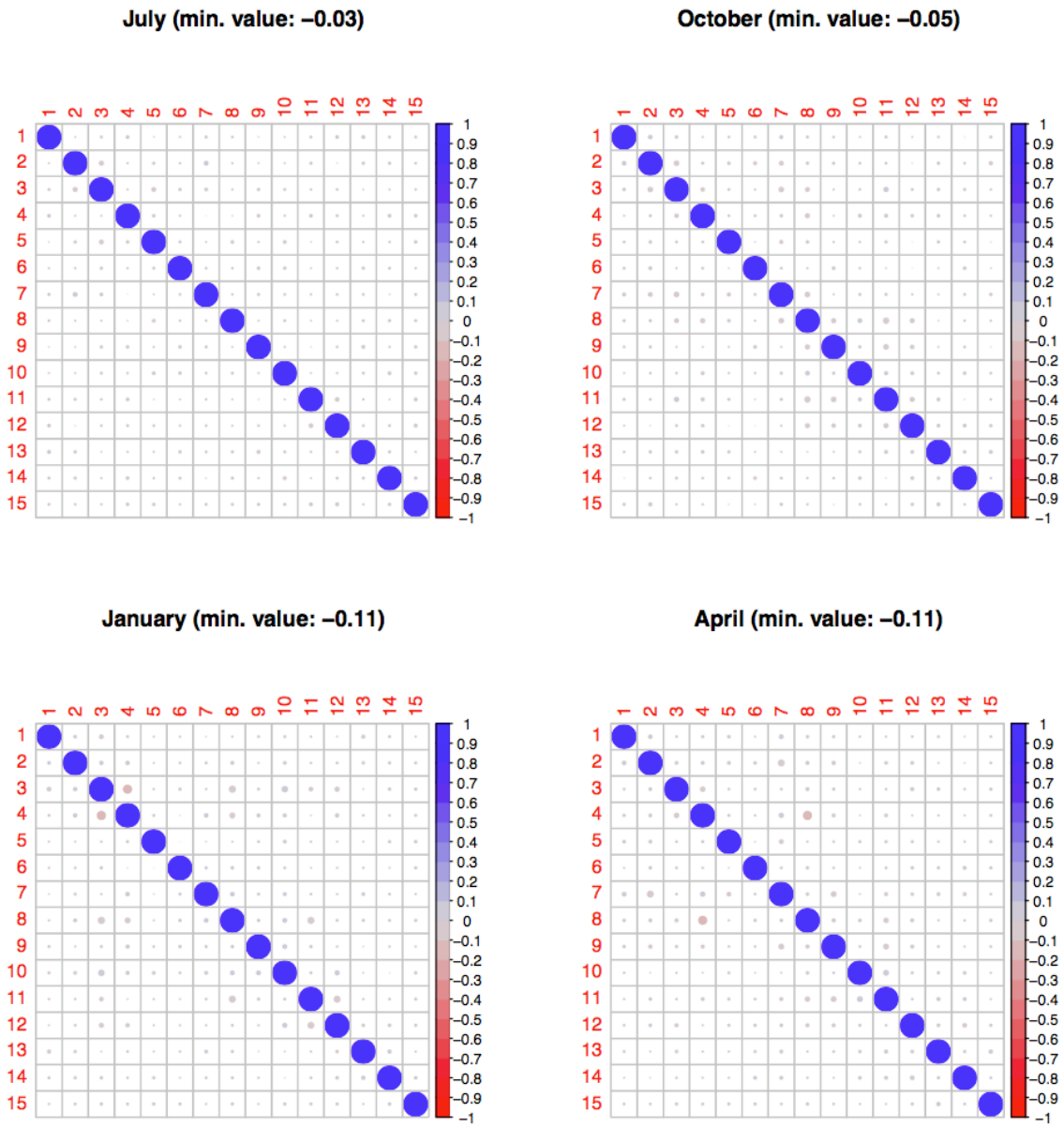

Figure S7. Correlations (between -1 and 1 or -100% and 100%) of posterior emissions between 15 different regions for the mid-month of each season. The anti-correlations in posterior emissions are only less than 20%. The numbers on the axis represent the region number, and the values in the parentheses represent the largest anti-correlation (negative value) for the given month.

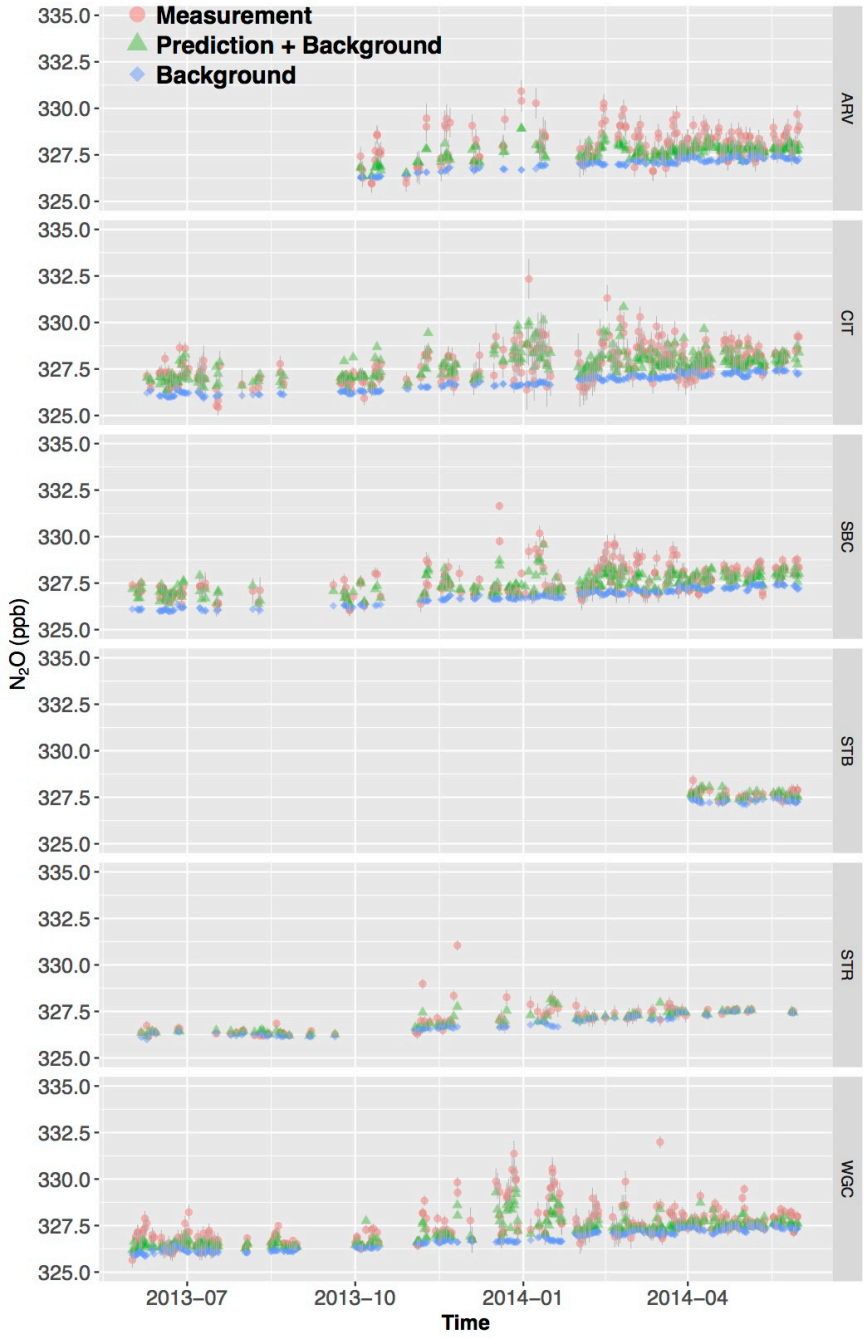

Figure S8. Comparison of measured and predicted N_2O mole fractions and estimated N_2O background by site. For measurements, we first calculate the bias corrected local mole fractions (i.e., total mole fraction - background - $\widehat{\boldsymbol{D}}$ where $\widehat{\boldsymbol{D}}$ is the optimized bias correction from Equation 1 in the main text) because we use local mole fractions in the inversion (to be compared with predictions). Then the estimated background is added to both measured (local) and predicted mole fractions for comparison. The error bar (in grey) represents the model-measurement uncertainty (i.e., median values (1σ) from Figure 2 in the main text).

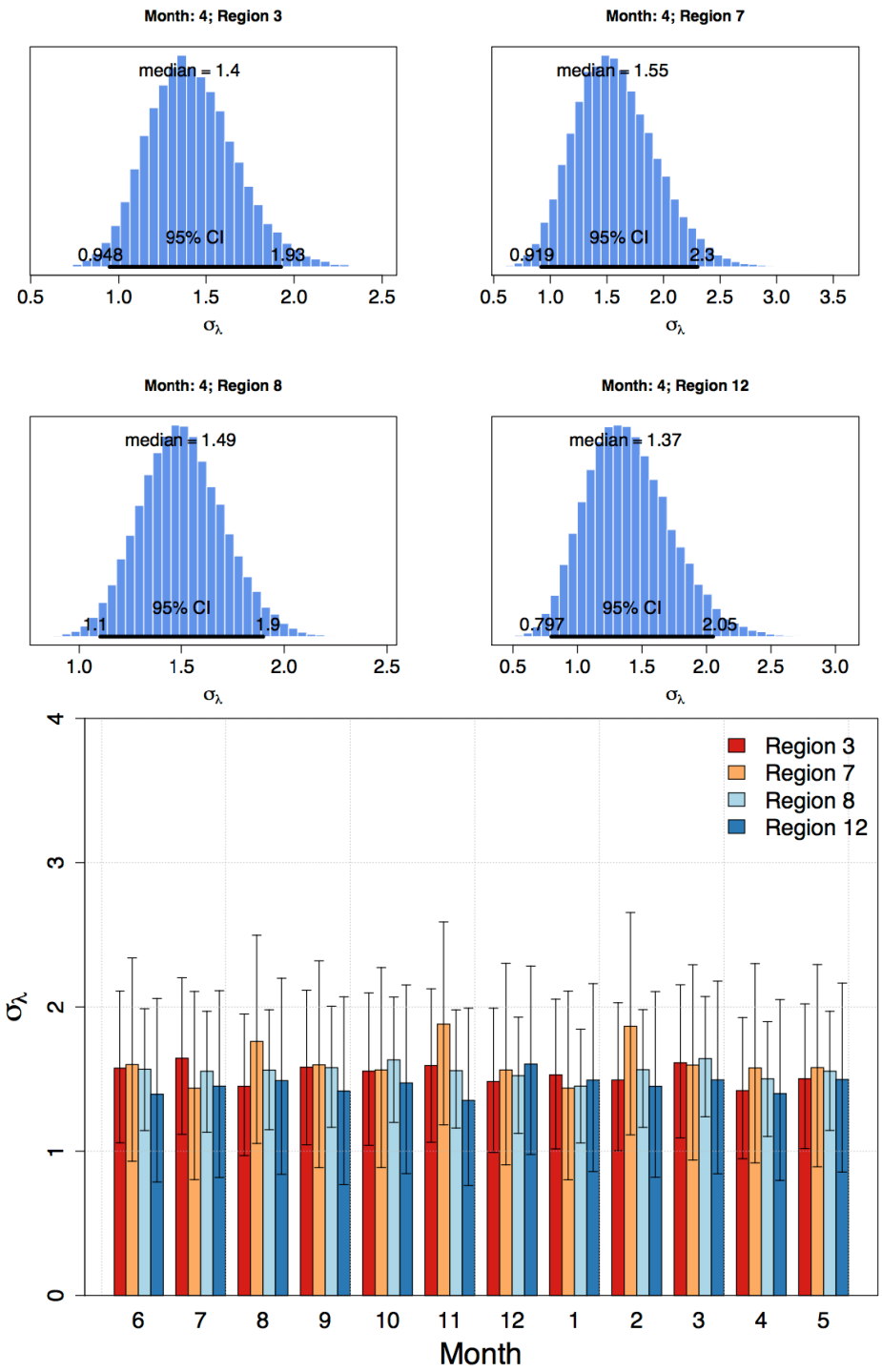

Figure S9. Example (April result) posterior distribution (histogram) of the N_2O prior uncertainty (i.e., $\boldsymbol{\sigma_\lambda}$) for the major regions (Regions 3, 7, 8 and 12) based on the 50000 samples (upper panel) and region-average posterior $\boldsymbol{\sigma_\lambda}$ for the major regions by month (lower panel). Although the posterior $\boldsymbol{\sigma_\lambda}$ estimates for individual 0.3° pixels within each region vary, here we present the region average value.

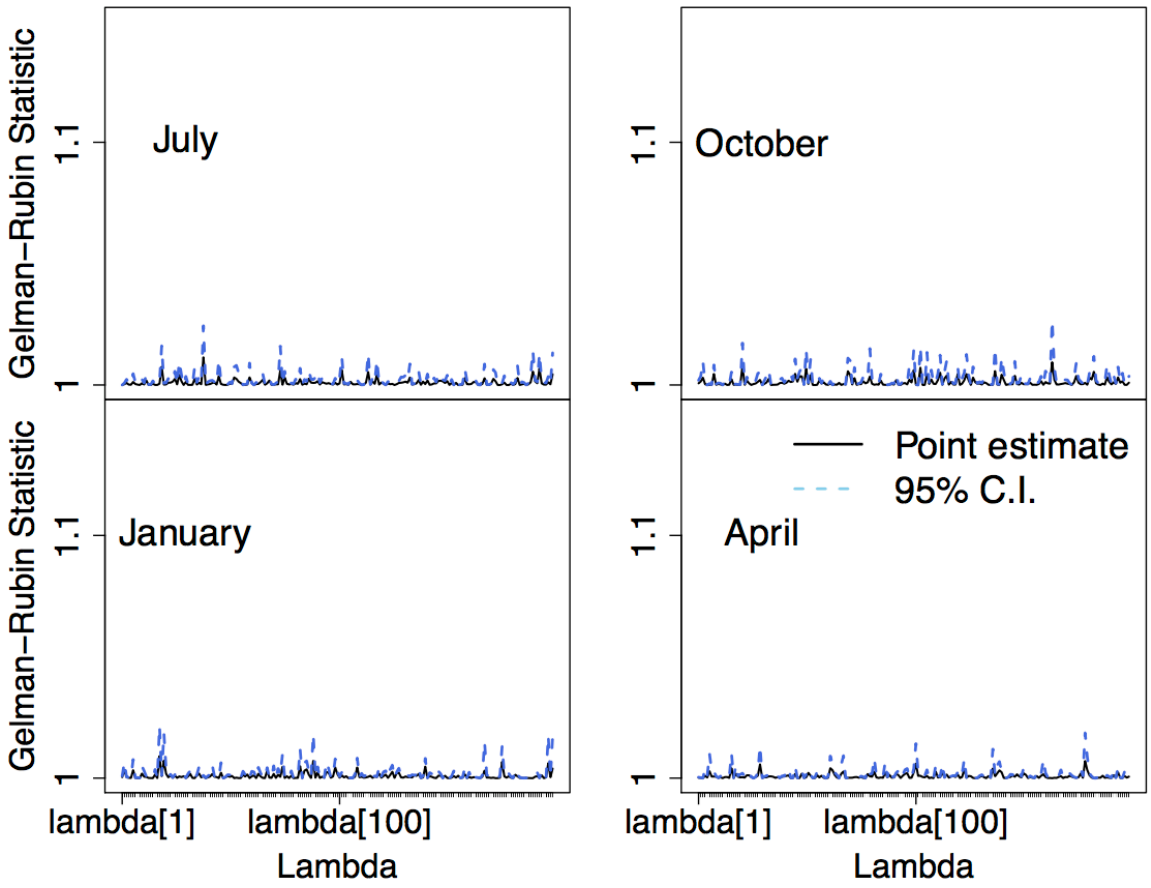

Figure S10. Gelman-Rubin statistics for the scaling factor vector (**λ**) for the mid-month of each season. The x-axis shows each element of the 197 scaling factors. For all scaling factors, the Gelman-Rubin statistics are close to one, suggesting the chains have successfully converged.

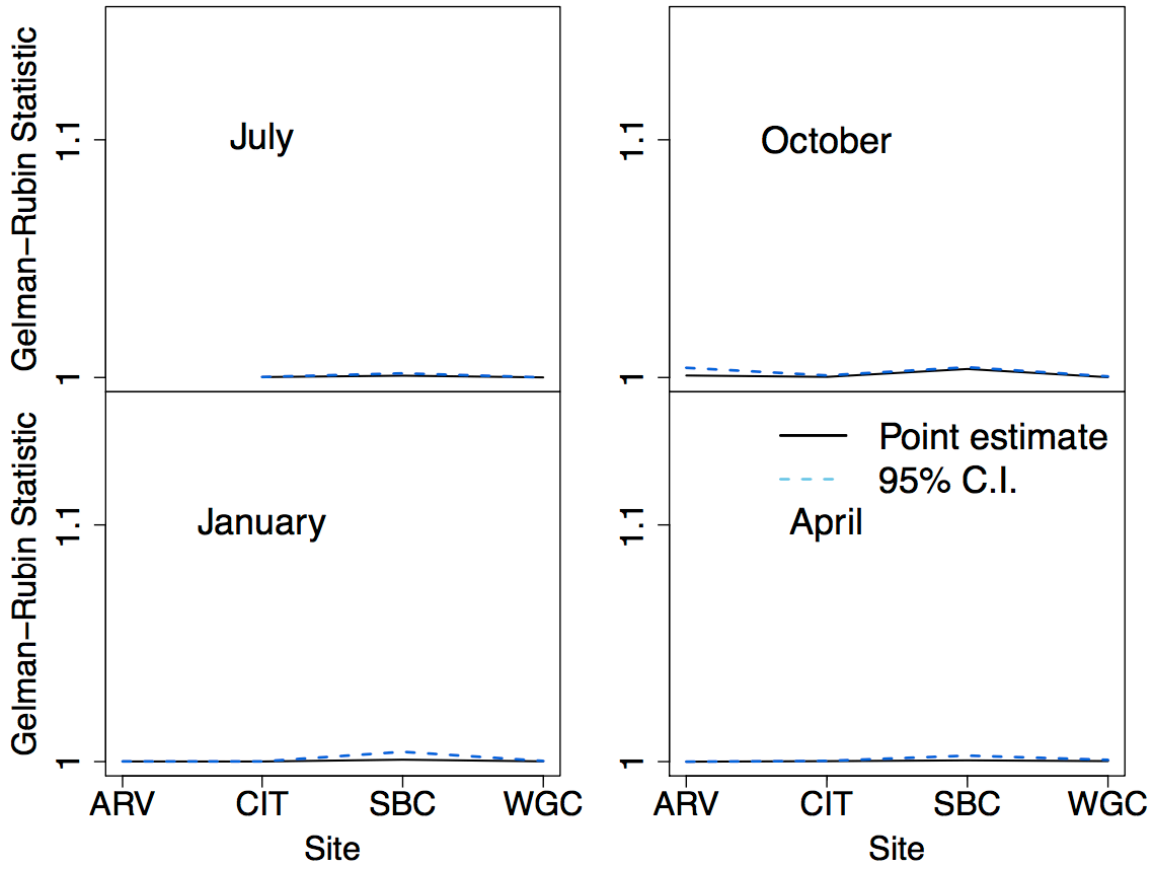

Figure S11. Gelman-Rubin statistics for σ_R at four sites in the Central Valley and SoCAB for the mid-month of each season. For all sites, the Gelman-Rubin statistics are close to one, suggesting the chains have successfully converged.

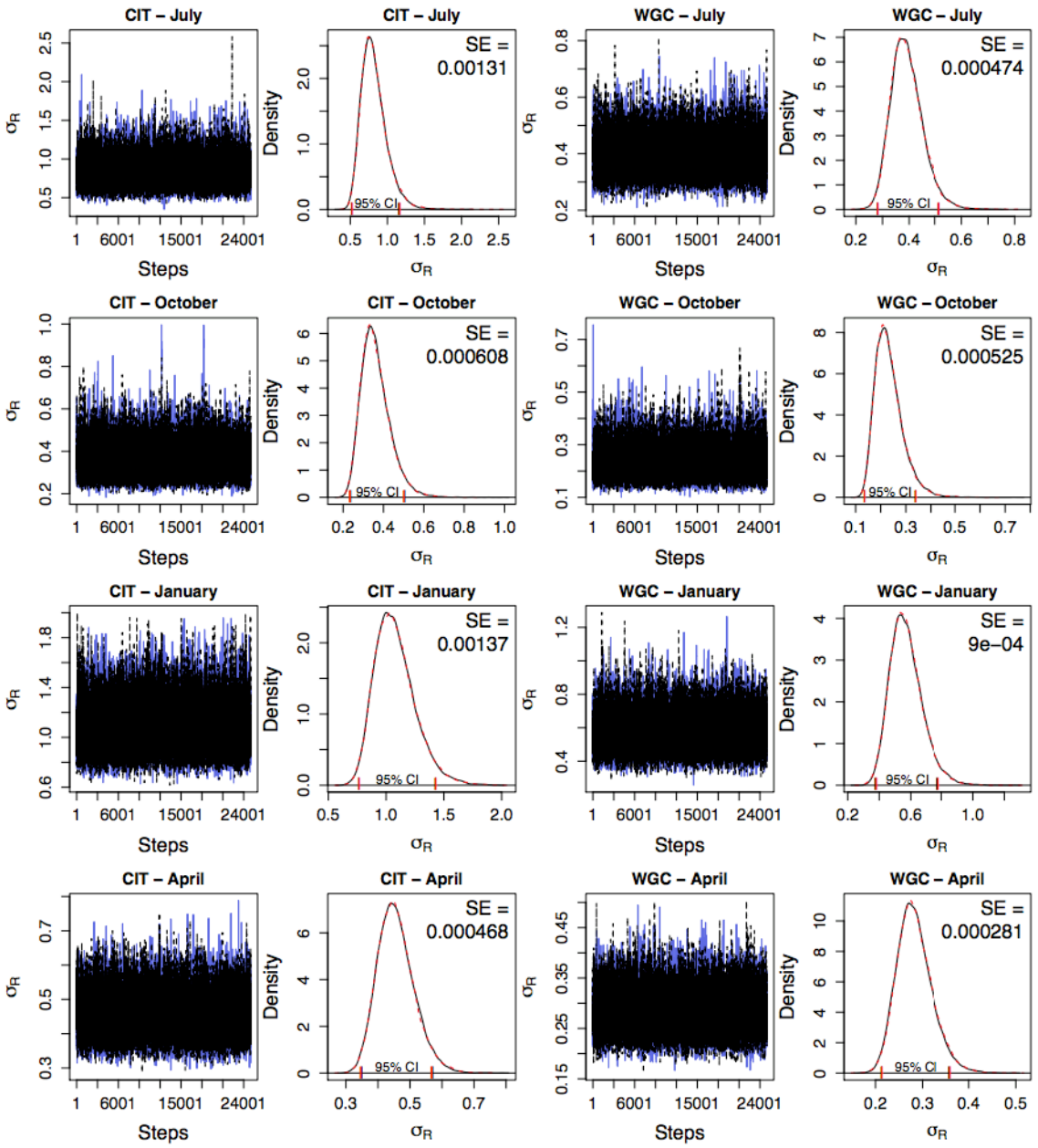

Figure S12. Example trace and density plots for the model-measurement mismatch uncertainty parameter ($\boldsymbol{\sigma_R}$) at WGC and CIT (for mid-month of each season). The two colors in each plot represent the two different chains. In the trace plots (first and third columns), the two chains are smooth and overlap each other suggesting the chains have converged well. For the posterior density (second and fourth columns), the two densities from two different chains overlap very well, suggesting the two chains are representative of the posterior distribution. In the density plots, we also include the standard error (SE) for the MCMC estimates using the parameter scale (here in units of ppb for $\boldsymbol{\sigma_R}$). The SE values are very small compared to the mean estimates.

References for Supplemental Information

See the references in the main text.